CLAIMS

What is claimed is:

1. A program storage device readable by a machine, tangibly embodying a program of instructions executable by the machine to perform method steps for constructing predictive models that can be used to make predictions even when the values of some or all inputs are missing or are otherwise unknown, the method steps comprising:

(1) presenting a collection of training data comprising examples of input values that are available to the model together with the corresponding desired output value(s) that the model is intended to predict;

and

(2) generating a plurality of subordinate models, that together comprise an overall model, in such a way that:

a) each subordinate model has an associated set of application conditions that must be satisfied in order to apply the subordinate model when making predictions, the application conditions comprising

i) tests for missing values for all, some, or none of the inputs,

and

ii) tests on the values of all, some, or none of the inputs that are applicable when the values of the inputs mentioned in the tests have known values;

and

b) for at least one subordinate model, the training cases used in the construction of that subordinate model include some cases that indirectly satisfy the application conditions in the sense that the application conditions are satisfied only after replacing one or more known data values in these training cases with missing values.

2. A device according to claim 1, wherein step (2) comprises generating a plurality of subordinate models with the further requirement that the plurality CANNOT be arranged into a decision-tree hierarchy in such a way that

(1) each branch of the tree corresponds to a test on the values of one or more data fields that can be satisfied only when those data fields have known values;

(2) each leaf of the tree corresponds to a subordinate model whose application conditions are defined by the conjunction of the tests along the branches that lead from the root node of the tree to the leaf node;

(3) the root node of the tree corresponds to a subordinate model whose application conditions consist of missing-value tests for the data fields mentioned in the tests associated with the tree branches that emanate from the root node;

and

(4) each interior node of the tree other than the root node corresponds to a subordinate model whose application conditions are defined by the conjunction of the tests along the branches that lead from the root node of the tree to the interior node, together with missing-value tests for the data fields mentioned in the tests associated with the tree branches that emanate from the interior node.